Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0826U001067

Особливі позначки: відкрита

Дата реєстрації: 16-04-2026

Статус: Запланована

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Батюк Тарас Миронович

2. Taras M. Batiuk

Кваліфікація:

Ідентифікатор ORCID ID: 0000-0001-5797-594X

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 124

Назва наукової спеціальності: Системний аналіз

Галузь / галузі знань:

Освітньо-наукова програма зі спеціальності: Системний аналіз

Дата захисту:

Спеціальність за освітою: Системний аналіз

Місце роботи здобувача:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): PhD 12869

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 28.29

Тема дисертації:

1. Інтелектуальна технологія кластеризації користувачів соціальних мереж на основі аналізу тональності повідомлень

2. Intelligent technology for clustering users of social networks based on the messages sentiment analysis

Реферат:

1. Дисертаційне дослідження присвячене розв'язанню актуального науково-прикладного завдання, що полягає у розробленні математичних методів та програмних засобів для кластеризації користувачів соціальних мереж на основі аналізу тональності та семантичного змісту їхніх текстових повідомлень. В умовах стрімкої цифровізації та зростання ролі онлайн-комунікацій соціальні мережі перетворилися на ключове середовище формування громадської думки, обміну інформацією та соціальної взаємодії. У зв'язку з цим ефективний аналіз великих обсягів неструктурованих текстових даних набуває особливого значення для дослідження поведінкових моделей користувачів, їхніх інтересів, емоційних станів та групових взаємозв'язків у віртуальному просторі. Такий підхід забезпечує більш точне та інформативне групування користувачів відповідно до їхніх емоційних, тематичних і поведінкових характеристик, а також дозволяє виявляти приховані структури у великих масивах соціальних даних. Наукова новизна отриманих результатів полягає у створенні інтелектуальної технології кластеризації користувачів соціальних мереж, яка інтегрує

сучасні методи аналізу текстів на основі глибокого навчання та алгоритми машинного навчання для автоматизованого виявлення груп користувачів із подібними емоційними та семантичними характеристиками. Вперше запропоновано комплексну інтеграцію моделей аналізу тональності та семантики з алгоритмом K-means у єдину технологічну схему, що дозволяє підвищити точність і стійкість результатів кластеризації. Окрім того, удосконалено методику попереднього опрацювання та нормалізації текстових даних, що сприяє підвищенню якості їх подальшого аналізу та інтерпретації. Практична значущість отриманих результатів полягає у можливості використання розроблених програмних засобів для автоматизованого аналізу соціальних даних з метою виявлення груп користувачів із подібними інтересами, емоційними станами та інформаційними потребами. Запропонована технологія може бути застосована в маркетингових дослідженнях, соціологічному аналізі, рекомендаційних системах, а також у діяльності державних органів і бізнес-структур для аналізу суспільних настроїв, оптимізації інформаційної політики та підвищення ефективності взаємодії з аудиторією в цифровому середовищі. У першому розділі здійснено комплексний огляд сучасних методів побудови технологій аналізу тональності, кластеризації повідомлень та авторизації й автентифікації користувачів. Проаналізовано ключові підходи до опрацювання текстових даних, зокрема методи семантичного та емоційного аналізу на основі глибинних нейронних мереж із акцентом на LSTM-архітектурах. Другий розділ дослідження присвячено концептуальному проектуванню інтелектуальної технології для аналізу тональності та кластеризації користувачів соціальних мереж. Розроблено концептуальну модель опрацювання даних, яка передбачає інтеграцію різнорідних джерел інформації та багатовимірних ознак, зокрема текстових векторів, LSTM-ознак емоційного стану та CNN-ознак візуального контенту. Третій розділ роботи присвячено навчанню та експериментальній апробації методів і програмних засобів інтелектуальної технології, спрямованої на аналіз тональності та подальшу кластеризацію користувачів соціальних мереж. Навчання моделей здійснювалося на попередньо підготовлених вибірках даних, сформованих на основі текстових повідомлень із соціальних мереж, з урахуванням процедур нормалізації та очищення даних. Для оцінювання якості побудованих моделей проведено порівняльний аналіз із використанням показників точності та повноти. Отримані експериментальні результати підтвердили ефективність поєднання LSTM, S-BERT та XGBoost для аналізу емоційного стану та семантичних характеристик текстів. Четвертий розділ спрямовано на практичну реалізацію та експериментальне навчання інтелектуальної технології кластеризації користувачів соціальних мереж на основі комплексного аналізу текстових і візуальних даних. У межах цього етапу результати аналізу тональності текстових повідомлень, а також ознаки, виділені з фотографій користувачів із використанням згорткових нейронних мереж CNN, інтегруються з вихідними даними алгоритму XGBoost для формування узагальненого вектора характеристик користувача. Подальший аналіз здійснюється із застосуванням алгоритму K-means, який виконує групування користувачів на основі багатовимірних векторів ознак, сформованих за результатами семантичного аналізу текстів, аналізу тональності та візуальних характеристик. Таким чином, розроблений підхід не лише підвищує аналітичну точність технології, але й створює інтерактивне середовище соціальної взаємодії, у якому користувачі можуть досліджувати власні вподобання, знаходити спільноти за інтересами та формувати персоналізований цифровий досвід.

2. The dissertation research is devoted to solving the current scientific and applied problem, which consists in developing mathematical methods and software tools for clustering social network users based on the analysis of the tone and semantic content of their text messages. In the conditions of rapid digitalization and the growth of the role of online communications, social networks have become a key environment for the formation of public opinion, information exchange and social interaction. In this regard, the effective analysis of large volumes of unstructured text data is of particular importance for studying the behavioral patterns of users, their interests, emotional states and group relationships in virtual space. This approach provides a more accurate and informative grouping of users according to their emotional, thematic and behavioral characteristics, and also allows you to identify hidden structures in large arrays of social data. The scientific novelty of the results obtained lies in the creation of an intelligent technology for clustering social network users, which integrates modern methods of text analysis based on deep learning and machine learning algorithms for automated detection of user groups with

similar emotional and semantic characteristics. For the first time, a comprehensive integration of tone and semantic analysis models with the K-means algorithm into a single technological scheme has been proposed, which allows to increase the accuracy and stability of clustering results. In addition, the method of pre-processing and normalization of text data has been improved, which contributes to improving the quality of their further analysis and interpretation. The practical significance of the results obtained lies in the possibility of using the developed software tools for automated analysis of social data in order to detect groups of users with similar interests, emotional states and information needs. The proposed technology can be applied in marketing research, sociological analysis, recommendation systems, as well as in the activities of government agencies and business structures to analyze public sentiment, optimize information policy and increase the efficiency of interaction with the audience in the digital environment. The first chapter provides a comprehensive review of modern methods for building technologies for sentiment analysis, message clustering, and user authorization and authentication. Key approaches to text data processing are analyzed, in particular, methods of semantic and emotional analysis based on deep neural networks with an emphasis on LSTM architectures. The second chapter of the study is devoted to the conceptual design of intelligent technology for sentiment analysis and clustering of social network users. A conceptual data processing model has been developed, which involves the integration of diverse sources of information and multidimensional features, in particular text vectors, LSTM features of emotional state and CNN features of visual content. The third chapter of the work is devoted to the training and experimental testing of methods and software tools of intelligent technology aimed at analyzing the tone and further clustering of users of social networks. The training of models was carried out on pre-prepared data samples formed on the basis of text messages from social networks, taking into account the procedures of normalization and data cleaning. To assess the quality of the constructed models, a comparative analysis was conducted using accuracy and completeness indicators. The experimental results obtained confirmed the effectiveness of the combination of LSTM, S-BERT and XGBoost for analyzing the emotional state and semantic characteristics of texts. The fourth chapter is aimed at the practical implementation and experimental training of the intelligent technology of clustering users of social networks based on the complex analysis of text and visual data. Within the framework of this stage, the results of the analysis of the tone of text messages, as well as features extracted from photos of users using convolutional neural networks CNN, are integrated with the output data of the XGBoost algorithm to form a generalized vector of user characteristics. Further analysis is carried out using the K-means algorithm, which performs user grouping based on multidimensional feature vectors formed based on the results of semantic text analysis, sentiment analysis, and visual characteristics. Thus, the developed approach not only increases the analytical accuracy of the technology, but also creates an interactive environment for social interaction in which users can explore their own preferences, find communities of interests, and create a personalized digital experience.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Теоретичне узагальнення і вирішення важливої наукової проблеми

Публікації:

- Батюк Т. М., Висоцька В. А. Розробка інтелектуальної системи підтримки соціалізації користувача за подібністю інтересів. Сучасний стан наукових досліджень та технологій в промисловості. 2022. № 1 (19). С. 13–26. DOI: https://doi.org/10.30837/ITSSI.2022.19.013
- Батюк Т. М., Висоцька В. А. Інформаційна підтримка процесів соціалізації особистості на основі спільних інтересів. Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі. 2022. Вип. 11. С. 56–86. DOI: https://doi.org/10.23939/sisn2022.11.056

- Батюк Т. М., Досин Д. Г. Імплементація інтелектуальної системи аналізу тональності та кластеризації публікацій у соціальній мережі Twitter. Сучасний стан наукових досліджень та технологій в промисловості. 2023. № 1 (23). С. 25–44. DOI: https://doi.org/10.30837/ITSSI.2023.23.025
- Батюк Т. М., Досин Д. Г. Інтелектуальна система кластеризації користувачів соціальних мереж на основі аналізу тональності даних. Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі. 2023. № 13. С. 121–138. DOI: https://doi.org/10.23939/sisn2023.13.121
- Батюк Т. М., Досин Д. Г. Реалізація надійної та ефективної автентифікації в інтелектуальних системах шляхом використання методів візуальної біометрики. Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі. 2024. № 15. С. 23–42. DOI: https://doi.org/10.23939/sisn2024.15.023
- Batiuk T., Dosyn D. Intellectual analysis of textual data in social networks using BERT and XGBoost. Вісник Національного університету "Львівська політехніка". Серія: Інформаційні системи та мережі. 2025. Вип. 17. С. 44–60. DOI: https://doi.org/10.23939/sisn2025.17.044
- Батюк Т. М., Висоцька В. А. Технологія соціалізації особистостей за спільними інтересами на основі методів машинного навчання та SEO-технологій. Радіоелектроніка, інформатика, управління. 2022. № 2 (61). С. 53–68. DOI: https://doi.org/10.15588/1607-3274-2022-2-6
- Batiuk T., Dosyn D. Realization of the decision-making support system for Twitter users' publications analysis. Радіоелектроніка, інформатика, управління. 2024. № 1 (68). С. 175–187. DOI: https://doi.org/10.15588/1607-3274-2024-1-16
- Batiuk T., Dosyn D. Optimizing authentication security in intelligent systems through visual biometrics for enhanced efficiency. Радіоелектроніка, інформатика, управління. 2024. № 3 (70). С. 54–66. DOI: https://doi.org/10.15588/1607-3274-2024-3-6
- Батюк Т. М., Досин Д. Г. Обробка текстових даних соціальних медіа на природній мові за допомогою BERT та XGBoost. Радіоелектроніка, інформатика, управління. 2025. № 2 (73). С. 154–167. DOI: https://doi.org/10.15588/1607-3274-2025-2-14

Наукова (науково-технічна) продукція: технології

Соціально-економічна спрямованість: забезпечення промисловості чи населення новим видом інформаційно-комунікаційних послуг

Охоронні документи на ОПІВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами:

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Досин Дмитро Григорович
2. Dmytro G. Dosyn

Кваліфікація: д. т. н., старший науковий співробітник, 05.13.06

Ідентифікатор ORCID ID: 0000-0003-4040-4467

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Безсонов Олександр Олександрович

2. Oleksandr O. Bezsonov

Кваліфікація: д.т.н., професор, 05.13.23

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи: Харківський національний університет радіоелектроніки

Код за ЄДРПОУ: 02071197

Місцезнаходження: проспект Науки, Харків, Харківський р-н., 61166, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Субботін Сергій Олександрович

2. Sergei Subbotin

Кваліфікація: д.т.н., професор, 05.13.23

Ідентифікатор ORCID ID: 0000-0001-5814-8268

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Запорізька політехніка"

Код за ЄДРПОУ: 02070849

Місцезнаходження: вул. Жуковського, Запоріжжя, Запорізький р-н., 69063, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Лозинська Ольга Володимирівна

2. Olha V. Lozynska

Кваліфікація: к. т. н., доц., 10.02.21

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Назаркевич Марія Андріївна

2. Mariia A. Nazarkevych

Кваліфікація: д.т.н., професор, 21.05.01

Ідентифікатор ORCID ID: 0000-0002-6528-9867

Додаткова інформація:

Повне найменування юридичної особи: Національний університет "Львівська політехніка"

Код за ЄДРПОУ: 02071010

Місцезнаходження: вул. Степана Бандери, Львів, 79013, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради	Пасічник Володимир Володимирович
Власне Прізвище Ім'я По-батькові головуючого на засіданні	Пасічник Володимир Володимирович
Відповідальний за підготовку облікових документів	Лучкевич Михайло Михайлович
Реєстратор	Юрченко Тетяна Анатоліївна

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності

Юрченко Тетяна Анатоліївна